

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan komunikasi telah mengalami kemajuan yang sangat pesat dalam beberapa tahun terakhir, salah satu hasil dari perkembangan ini adalah munculnya media sosial yang telah mengubah cara kita berkomunikasi dalam masyarakat. Media sosial merupakan platform yang memungkinkan pengguna untuk berinteraksi, berbagi informasi, dan membentuk ikatan sosial secara virtual. Menurut (Siregar, 2022) media sosial memiliki peran penting dalam berbagai bidang seperti komunikasi pemasaran, komunikasi politik, dan komunikasi dalam sistem pembelajaran, komunikasi mengenai opini publik.

X sebagai salah satu media sosial dengan jumlah pengguna terbanyak, memainkan peran penting dalam perubahan ini, X adalah jejaring sosial yang memungkinkan pengguna untuk mengirim pesan singkat atau *tweet* dengan batasan 140 karakter. Platform ini memungkinkan pengguna untuk membagikan foto, video, serta berinteraksi dengan pengguna lainnya secara real-time. Menurut (Handayani, 2021) media sosial seperti X telah menjadi tempat yang penting bagi pengguna untuk menyampaikan pendapat mereka tentang berbagai topik, termasuk mengenai politik, kebijakan publik dan tokoh politik. Analisis sentimen pada media sosial dapat memberikan wawasan tentang persepsi publik terhadap isu-isu tertentu.

Anies Baswedan adalah seorang tokoh politik yang diumumkan oleh partai Nasional Demokrat pada 3 Oktober 2022 sebagai bakal calon presiden Indonesia untuk tahun 2024. Sebagai figur politik yang menjadi pusat perhatian, Anies Baswedan mendapat banyak perhatian di media sosial. Menurut analisis yang dilakukan oleh Drone Emprit (Rahman, 2022) pada periode 3-10 Juli 2022, Anies Baswedan menjadi tokoh politik yang paling banyak dibicarakan di media sosial dengan 105.110 sebutan. Opini atau komentar mengenai Anies Baswedan dapat disampaikan oleh masyarakat melalui media sosial X dengan mengunggah cuitan

atau *tweet*. Data dari *tweet* ini dapat digunakan untuk mendapatkan gambaran positif dan negatif atau netral melalui analisis sentimen.

Dalam era digital saat ini, penting bagi tokoh politik dan pembuat kebijakan untuk memahami bagaimana mereka dan kebijakan mereka dilihat oleh publik. Namun dengan volume besar data yang tersedia di media sosial, ada tantangan dalam menyaring dan menganalisis opini yang relevan secara manual. Metode tradisional seperti survei dan wawancara mungkin tidak lagi cukup efektif untuk menangkap sentimen publik secara real-time dan dalam skala besar. Oleh karena itu, diperlukan metode yang lebih efisien dan akurat untuk menganalisis sentimen ini. Meskipun ada banyak penelitian tentang analisis sentimen, studi yang secara khusus menganalisis sentimen publik terhadap Anies Baswedan di media sosial X dengan menggunakan metode *machine learning* masih terbatas.

Analisis sentimen merupakan cabang dari *text mining* yang berfokus pada pengklasifikasian teks berdasarkan opini atau sentimen yang terkandung di dalamnya. (Idris, Mustofa and Salihi, 2023) menyebutkan bahwa analisis sentimen dapat melakukan ekstraksi pendapat, emosi, dan evaluasi tertulis seseorang tentang topik tertentu menggunakan teknik pemrosesan bahasa alami. Dalam konteks tokoh politik yaitu Anies Baswedan, analisis sentimen dapat membantu mengidentifikasi sentimen positif, negative atau netral yang disampaikan oleh pengguna media sosial mengenai kebijakan sosial atau politik atau juga terhadap tokoh politik tersebut.

Penggunaan metode *Support Vector Machine* (SVM) dan *Random Forest* untuk analisis sentimen telah terbukti efektif dalam berbagai penelitian *Support Vector Machine* berhasil mengklasifikasikan opini pengguna X dengan tingkat akurasi sebesar 93%. Sementara itu, (Idris, Mustofa and Salihi, 2023) menemukan bahwa *Support Vector Machine* mampu menghasilkan akurasi sebesar 98% dalam mengklasifikasikan sentimen pengguna aplikasi *Shopee*. Metode *Random Forest* juga telah digunakan dalam analisis sentimen dengan hasil yang memuaskan. Menurut (Afdhal *et al.*, 2022) *Random Forest* dapat menghasilkan kinerja yang baik dalam klasifikasi sentimen komentar di *YouTube* dengan tingkat akurasi yang tinggi.

Berdasarkan latar belakang tersebut, maka penulis melakukan penelitian dengan judul “ANALISIS SENTIMEN TERHADAP ANIES BASWEDAN DI MEDIA SOSIAL X DENGAN SUPPORT VECTOR MACHINE DAN RANDOM FOREST”. Penelitian ini bertujuan untuk memahami sentimen publik terhadap Anies Baswedan serta mengevaluasi efektivitas dua algoritma *machine learning* dalam menganalisis data dari media sosial X.

1.2 Perumusan Masalah

Berdasarkan latar belakang ini, masalah penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana mengembangkan analisis sentimen terhadap Anies Baswedan di Media Sosial X menggunakan algoritma *Support Vector Machine* (SVM) dan *Random Forest*?
2. Bagaimana hasil klasifikasi sentimen yang dihasilkan oleh algoritma *Support Vector Machine* (SVM) dan *Random Forest* berdasarkan evaluasi dengan *confusion matrix* serta bagaimana perbandingan performa antara kedua algoritma tersebut?

1.3 Batasan Masalah

Penelitian ini memiliki beberapa batasan masalah yang perlu diperhatikan untuk memastikan fokus penelitian tetap terarah dan hasil yang diperoleh relevan serta dapat dipertanggungjawabkan. Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini hanya menggunakan data dari Media Sosial X, data yang diambil terbatas pada komentar dan postingan publik yang berkaitan dengan Anies Baswedan, tidak ada data dari media sosial lainnya yang akan digunakan dalam analisis ini.
2. Analisis sentimen yang dilakukan dalam penelitian ini dibatasi pada teks berbahasa Indonesia. Teks dalam bahasa lain yang mungkin muncul di Media Sosial X tidak akan dimasukkan dalam analisis.
3. Penelitian ini hanya menggunakan dua metode analisis sentimen yaitu *Support*

Vector Machine (SVM) dan *Random Forest*. Metode analisis sentimen lainnya seperti *Naïve Bayes*, *K-Nearest Neighbor* (KNN), atau *Deep Learning* tidak akan digunakan dalam penelitian ini.

4. Analisis sentimen dalam penelitian ini terbatas pada klasifikasi tiga jenis sentimen yaitu positif, negatif, dan netral. Sentimen yang lebih kompleks seperti emosi (marah, senang, sedih, dll.) tidak akan dianalisis.
5. Proses *preprocessing data* mencakup tahapan-tahapan seperti labeling manual, *cleansing*, *case folding tokenizing*, *stopword removal* dan *stemming*. Batasan ini diperlukan untuk mengurangi kebisingan dalam data dan memastikan data yang dianalisis berada dalam format yang konsisten.
6. Asumsi yang digunakan dalam penelitian ini adalah bahwa data yang diambil dari Media Sosial X adalah representatif dari pendapat umum pengguna media sosial tersebut mengenai Anies Baswedan. Tidak ada verifikasi tambahan yang dilakukan untuk memastikan keaslian atau representativitas data.

Batasan-batasan ini dibuat untuk memastikan penelitian tetap fokus, terarah, dan selesai dalam waktu yang telah ditentukan. Selain itu batasan-batasan juga membantu mengelola kompleksitas penelitian dan memastikan bahwa hasilnya dapat dianalisis dan disimpulkan dengan benar.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Mengembangkan analisis sentimen terhadap Anies Baswedan di Media Sosial X menggunakan algoritma *Support Vector Machine* (SVM) dan *Random Forest*.
2. Menganalisis hasil klasifikasi sentimen dengan menggunakan *confusion matrix* serta membandingkan performa antara algoritma *Support Vector Machine* (SVM) dan *Random Forest*.

Dengan mencapai tujuan-tujuan tersebut, diharapkan penelitian ini akan memberikan wawasan yang lebih mendalam tentang pandangan publik tentang Anies Baswedan dan membantu dalam pembentukan praktik dan kebijakan yang lebih baik.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan berbagai manfaat yang spesifik, terukur, dan relevan berdasarkan tujuan penelitian yang telah dirumuskan. Berikut adalah manfaat yang diharapkan dari penelitian ini:

1. Penelitian ini akan memperkaya literatur tentang analisis sentimen di media sosial, khususnya yang berkaitan dengan tokoh politik Anies Baswedan. Hasil penelitian dapat menjadi referensi bagi penelitian-penelitian selanjutnya yang ingin mengkaji topik serupa atau mengembangkan metode analisis sentimen.
2. Penelitian ini memberikan kontribusi dalam metode analisis sentimen dengan membandingkan performa dua algoritma yang berbeda yaitu *Support Vector Machine* (SVM) dan *Random Forest*. Temuan ini dapat digunakan untuk meningkatkan metode analisis sentimen di berbagai bidang penelitian lainnya.
3. Hasil analisis sentimen yang diperoleh dari penelitian ini dapat memberikan wawasan penting bagi pembuat kebijakan tentang persepsi publik terhadap Anies Baswedan. Informasi ini dapat digunakan untuk memahami bagaimana kebijakan-kebijakan yang diambil oleh Anies Baswedan diterima oleh publik dan bagaimana respons masyarakat terhadap tindakan-tindakan politiknya. Dengan demikian, hasil penelitian ini dapat membantu pembuat kebijakan dalam merumuskan strategi yang lebih efektif dan responsif terhadap aspirasi masyarakat.
4. Dengan memahami sentimen publik terhadap Anies Baswedan, penelitian ini juga dapat berkontribusi dalam meningkatkan kesadaran masyarakat tentang pentingnya mengemukakan opini secara konstruktif di media sosial. Selain itu, penelitian ini dapat menjadi sarana edukasi bagi tokoh politik dan masyarakat tentang pentingnya mendengarkan dan memahami persepsi publik sebagai bagian dari proses demokrasi.
5. Penelitian ini diharapkan dapat membantu membantu tokoh politik untuk membantu memahami aspirasi masyarakat terhadap kebijakan, kekhawatiran masyarakat serta dapat meresponsnya dengan kebijakan yang lebih tepat dan relevan yang akan diambil atau dilakukan.

Dengan demikian, penelitian ini tidak hanya memberikan kontribusi untuk kemajuan ilmu pengetahuan dan metodologi, tetapi juga memiliki efek sosial dan praktis yang menguntungkan bagi berbagai pihak yang terkait.

1.6 Metodologi Penelitian

Penelitian ini menggunakan metode yang terstruktur untuk mengumpulkan dan menganalisis data terkait Anies Baswedan di Media Sosial X

1.6.1 Metode Pengumpulan Data

1. Sumber Data:

Data yang digunakan dalam penelitian ini berasal dari postingan publik di Media Sosial X yang mengandung kata kunci "Anies Baswedan". Data dikumpulkan menggunakan teknik *crawling*.

2. Alat dan Software:

Data dikumpulkan menggunakan *Google Collab* dan bahasa pemrograman *Python*. Teknik *crawling* dilakukan dengan menggunakan library dan *framework* yang tersedia di *Python*.

3. Proses Pengumpulan Data:

Proses pengumpulan data dilakukan dengan teknik *crawling*, yang melibatkan identifikasi dan pengumpulan data yang relevan berdasarkan kata kunci yang telah ditentukan. Teknik ini memungkinkan pengambilan data secara otomatis dan sistematis dari Media Sosial X.

1.6.2 Metode Analisis Data

1. *Preprocessing* Data:

Preprocessing Data adalah serangkaian langkah yang dilakukan untuk mempersiapkan data mentah agar dapat dianalisis secara efektif. Tujuannya adalah untuk membersihkan dan mengubah data mentah menjadi format yang sesuai dan konsisten, sehingga algoritma *machine learning* dapat mengolahnya dengan lebih baik. Dalam konteks analisis sentimen di media sosial *preprocessing* data sangat penting karena data

teks biasanya mengandung banyak kebisingan, seperti tanda baca, karakter khusus, dan berbagai bentuk variasi bahasa. Langkah-langkah *preprocessing* yang diterapkan pada data teks dalam penelitian ini meliputi:

a. *Labeling Manual*

Pada tahapan awal data teks akan diberikan sentimen sesuai dengan klasifikasinya, yaitu positif, negatif, atau netral, pelabelan dilakukan secara manual oleh peneliti yang membaca dan menginterpretasikan konten teks untuk menentukan sentimen yang sesuai. Proses ini, meskipun bersifat subjektif, penting untuk memastikan bahwa data pelatihan memiliki label yang akurat.

b. *Cleaning*

Proses ini melibatkan penghapusan karakter atau simbol yang tidak diperlukan, seperti tanda baca, emotikon, dan *HTML tags*. *Cleaning* bertujuan untuk menghilangkan elemen-elemen yang tidak relevan dan dapat mengganggu analisis data.

c. *Case Folding*

Semua teks diubah menjadi huruf kecil untuk memastikan konsistensi. Misalnya, kata "Anies" dan "anies" akan dianggap sama setelah *case folding*. Ini membantu dalam mengurangi variasi kata yang sebenarnya memiliki arti yang sama.

d. *Tokenizing*

Proses ini memecah teks menjadi unit-unit yang lebih kecil yang disebut token. *Tokenizing* memudahkan analisis dengan mengubah teks menjadi daftar kata yang dapat diolah oleh algoritma.

e. *Stopword Removal/Filtering*

Stopwords adalah kata-kata umum yang sering muncul dalam teks tetapi tidak memiliki makna signifikan dalam analisis, seperti "dan", "atau", "yang". Proses ini menghapus stopwords untuk meningkatkan fokus pada kata-kata yang lebih bermakna.

f. *Stemming*

Stemming adalah proses mengubah kata-kata menjadi bentuk dasarnya. Misalnya kata "berjalan", "berjalan-jalan", dan "berjalanlah" akan diubah menjadi "jalan". Ini membantu dalam konsolidasi variasi kata yang memiliki akar yang sama.

2. Implementasi Algoritma:

Algoritma yang digunakan dalam penelitian ini adalah *Support Vector Machine* (SVM) dan *Random Forest*. Implementasi dilakukan menggunakan library *Scikit-Learn* (*sklearn*) di *Python*.

3. Evaluasi Model:

Evaluasi performa model dilakukan menggunakan teknik *K-Fold Cross-Validation*. Metrik evaluasi yang digunakan meliputi akurasi, presisi, dan *recall*. Teknik ini memastikan bahwa hasil evaluasi model lebih andal dan tidak bias terhadap satu *subset* data tertentu.

4. Perbandingan Algoritma:

Hasil dari kedua algoritma (*Support Vector Machine* dan *Random Forest*) dibandingkan menggunakan teknik statistik dan visualisasi data. Perbandingan ini dilakukan untuk menentukan algoritma mana yang lebih efektif dalam mengklasifikasikan sentimen publik terkait Anies Baswedan di Media Sosial X.

1.7 Sistematika Penulisan

Sistematika penulisan dibuat untuk memberikan gambaran luas tentang penelitian yang dilakukan dan untuk memberikan penjelasan singkat tentang kerangka penulisan.

BAB I: PENDAHULUAN

Bab ini berisi latar belakang masalah, perumusan masalah, batasan masalah, maksud dan tujuan, manfaat penelitian, metode penelitian, dan sistematika penulisan. Bagian ini memberikan gambaran umum mengenai konteks dan fokus penelitian serta langkah-langkah metodologis yang akan diambil.

BAB II: TINJAUAN PUSTAKA

Bab ini akan membahas teori-teori yang relevan dengan topik penelitian, termasuk Anies Baswedan, analisis sentimen, metode *Support Vector Machine* (SVM), dan *Random Forest*. Selain itu, bab ini juga akan menyajikan hasil penelitian sebelumnya yang berkaitan dengan topik ini untuk memberikan dasar teoritis yang kuat bagi penelitian.

BAB III: METODE PENELITIAN

Bab ini mencakup penjelasan tentang metode penelitian yang digunakan dalam penelitian ini. Ini mencakup metode pengumpulan data, teknik *preprocessing* data, implementasi algoritma *Support Vector Machine* dan *Random Forest*, serta teknik evaluasi model. Bab ini juga akan menjelaskan bagaimana data dianalisis secara keseluruhan.

BAB IV: HASIL DAN PEMBAHASAN

Bab ini berisi implementasi, pengujian, dan analisis hasil penelitian. Tujuan dari bab ini adalah untuk menentukan seberapa efektif metode *Support Vector Machine* dan *Random Forest* dalam mengklasifikasikan sentimen publik terkait Anies Baswedan di Media Sosial X. Selain itu, bab ini akan membahas perbandingan performa kedua algoritma dan interpretasi hasil yang diperoleh.

BAB V: PENUTUP

Bab ini menyajikan kesimpulan dari hasil penelitian dan memberikan rekomendasi untuk penelitian masa depan. Ini mencakup peningkatan algoritma klasifikasi, penambahan fitur pada sistem analisis sentimen, dan saran untuk pengembangan lebih lanjut mengenai topik sentimen Anies Baswedan. Bab ini juga akan memberikan refleksi tentang keterbatasan penelitian dan implikasi praktis dari temuan yang diperoleh.

Dengan sistematika penulisan ini, diharapkan dapat memberikan alur yang jelas dan terstruktur dalam penyampaian penelitian sehingga pembaca dapat memahami konteks, metodologi, hasil, dan kontribusi penelitian ini dengan baik.